IOP Conf. Series: Materials Science and Engineering **394** (2018) 022031 doi:10.1088/1757-899X/394/2/022031

Bioinformatics Analysis of the Genes *CYP79B2* in Cabbage (*Brassica oleracea* var. *capitata*)

Bo Sun, Min Jiang, Qiao Yuan, Fen Zhang, and Haoru Tang*

College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China.

*Corresponding author e-mail: htang@sicau.edu.cn

Abstract. CYP79B2 is an important cytochrome P450 monooxygenases in glucosinolate biosynthesis. Here, the *Brassica* oleracea var. *capitata CYP79B2* (*BocCYP79B2*) gene sequences were obtained from *Brassica* database (BRAD), and preformed for bioinformatics analysis. The *BocCYP79B2.1*, *BocCYP79B2.2* and *BocCYP79B2.3* genes mapped to chromosomes 1, 3 and 7, and contains an open reading frame of 1,623 bp, 1,557 bp and 1,626 bp that encodes a 540, 518, 541 amino acid protein, respectively. Subcellular localization predicted all *BocCYP79B2* genes were in the chloroplast. The conserved domain of the BocCYP79B2 protein is PLN02971. Homology analysis indicates that the CYP79B2 protein is apparently conserved during plant evolution. The findings of the present study provide a molecular basis for the elucidation of *CYP79B2* gene function in cabbage.

1. Introduction

Cabbage (*Brassica oleracea* var. *capitata*) is a member of the Brassicaceae family that is widely distributed in the world. In China, cabbage is an important vegetable crop, and consumed considerable every years. Cabbage is generally grown for its leafy head as common edible part, which are crispy, tender, and tasty [1]. Besides its good flavor, cabbage is also a rich source of nutrients, antioxidants, and anticarcinogenic compounds, including carbohydrates, vitamin C, carotenoids, and glucosinolates [1-2].

Glucosinolates are a group of sulfur- and nitrogen-containing secondary metabolites that are mainly found in the order of Brassicales and related groups of dicotyledonous angiosperms [3-4]. Glucosinolates and the hydrolytic myrosinase (β-thiosglucoside glucohydrolase) are stored separately under normal situations, but they come into contact with each other when tissues are damaged, and then the glucosinolates are hydrolyzed into several degradation products, such as isothiocyanates and nitriles [5]. Glucosinolates and their degradation products have diverse biological functions, which contribute to human health, as well as the taste and odor of cruciferous crops. The anticancer activity of isothiocyanates has been widely studied, and the mechanism involved has been elucidated [6].

Glucosinolate metabolism in plants is modulated by numerous biotic and abiotic factors, and the regulatory network of glucosinolate metabolism has been well elucidated in *Arabidopsis* [6]. CYP79B2 belongs cytochrome P450 monooxygenases (cytochromes P450) of the CYP79 family. CYP79B2 together with the homolog CYP79B3 catalyze the conversion of tryptophan to indole-3-acetaldoxime (IAOx) [7]. The gene encoding the CYP79B2 protein has been isolated in *Arabidopsis thaliana* and Chinese cabbage [7-8]. To date, research studies on *CYP79B2* in cabbage are limited. In

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

the present study, the *CYP79B2* gene sequence of cabbage was obtained from web database, and then sequence analysis of the *CYP79B2* gene were analyzed. The present study aimed to establish the foundation for further studies on the molecular mechanism of *CYP79B2* in cabbage.

2. Materials and methods

2.1. Sequence Obtain of the BocCYP79B2 Gene

The genomic DNA and mRNA sequences of *CYP79B2* gene of cabbage were downloaded and obtained from The *Brassica* database (BRAD) (http://brassicadb.org), and then used to subsequent bioinformatic analysis.

2.2. Bioinformatics Analysis of the BocCYP79B2 Gene

The amino acid sequence, protein molecular weight, isoelectric point, stability index, and hydrophobicity of the *BocCYP79B2* gene were analyzed and predicted by ExPASy (http://web.expasy.org) and NCBI (https://www.ncbi.nlm.nih.gov/). The prediction of protein secondary structure was done using DNAStar software. Subcellular localization was predicted by WoLF PSORT (http://www.genscript.com/wolf-psort.html). The conserved domain were predicted by NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Phylogenetic tree analysis of the CYP79B2 proteins was executed in MEGA 6.0 using the neighbor-joining (NJ) method.

3. Results

3.1. Analysis on genomic organization

The *Brassica* database (BRAD) was used to analyze the chromosomal localization and genomic organization of *BocCYP79B2*. There are three genes of *CYP79B2* in cabbage chromosomes, *BocCYP79B2.1*, *BocCYP79B2.2* and *BocCYP79B2.3*, and the gene IDs in BRAD are Bol028852, Bol032767 and Bol018585, respectively. The *BocCYP79B2.1* gene was mapped to chromosomes 1 and has 2 exons and 1 intron, the *BocCYP79B2.2* gene was mapped to chromosomes 3 and has 2 exons and 1 intron, and the *BocCYP79B2.3* gene was mapped to chromosomes 7 and has 2 exons and 1 intron (Fig. 1).

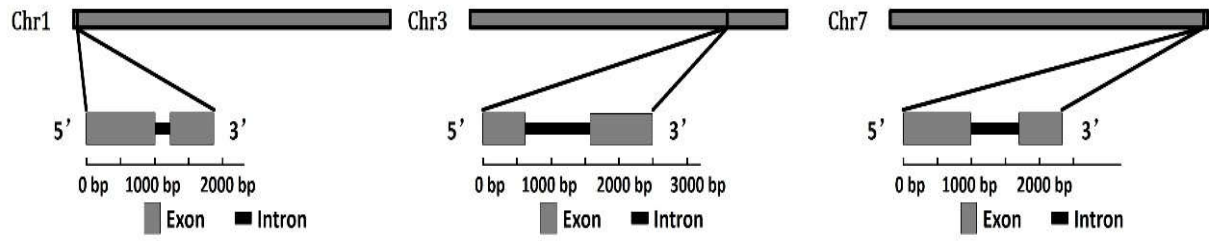


Fig. 1 Chromosomal location and genomic structure of *BocCYP79B2*.

3.2. Protein physical and chemical properties analysis

Sequence analysis indicated that the *BocCYP79B2.1*, *BocCYP79B2.2* and *BocCYP79B2.3* gene contained a 1,623 bp, 1,557 bp and 1,626 bp open reading frame (ORF), which encoded a 540, 518, 541 amino acids protein with a calculated molecular mass of 61.13 kD, 58.78 kD and 61.31 kD, and an isoelectric point (pI) of 8.81, 9 and 8.8, respectively. The amino acid types and proportions of the *Boc CYP79B2* genes were shown in Figure 2, the highest number in each gene is Leucine (Leu), the lowest number of amino acid in *BocCYP79B2.1* and *BocCYP79B2.2* is Tryptophan (Trp), the lowest number in *BocCYP79B2.3* is Cysteine (Cys). The predicted formula of *BocCYP79B2.1*, *BocCYP79B2.2* and *BocCYP79B2.3* were $C_{2736}H_{4346}N_{740}O_{779}S_{33}$, $C_{2643}H_{4217}N_{713}O_{732}S_{34}$ and $C_{2744}H_{4362}N_{740}O_{779}S_{35}$, respectively. Their total average hydrophilicity index were -0.196, -0.133 and -0.188, liposoluble index were 85.78, 91.7 and 87.08, and instability index in solution were 35.53,

36.95 and 34.73, respectively. *BocCYP79B2.3* has two transmembrane structure, whereas *BocCYP79B2.1* and *BocCYP79B2.2* have none (Fig. 3).

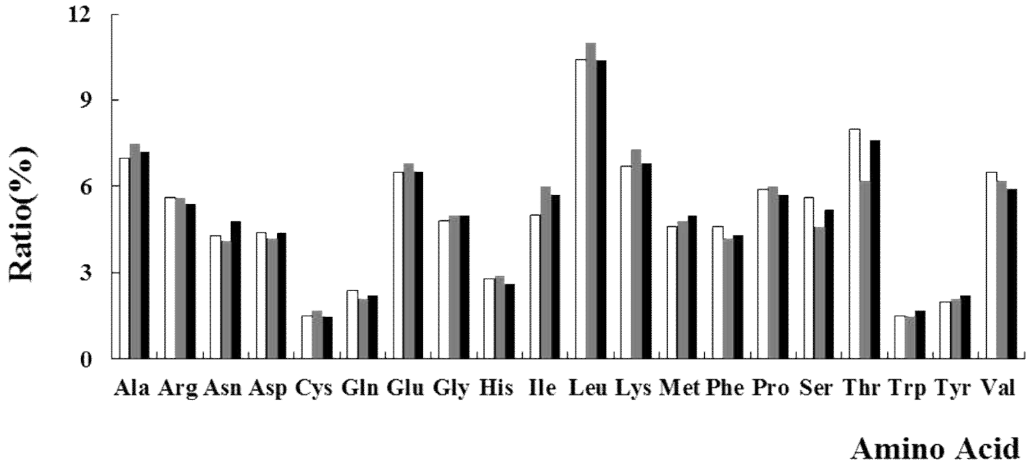
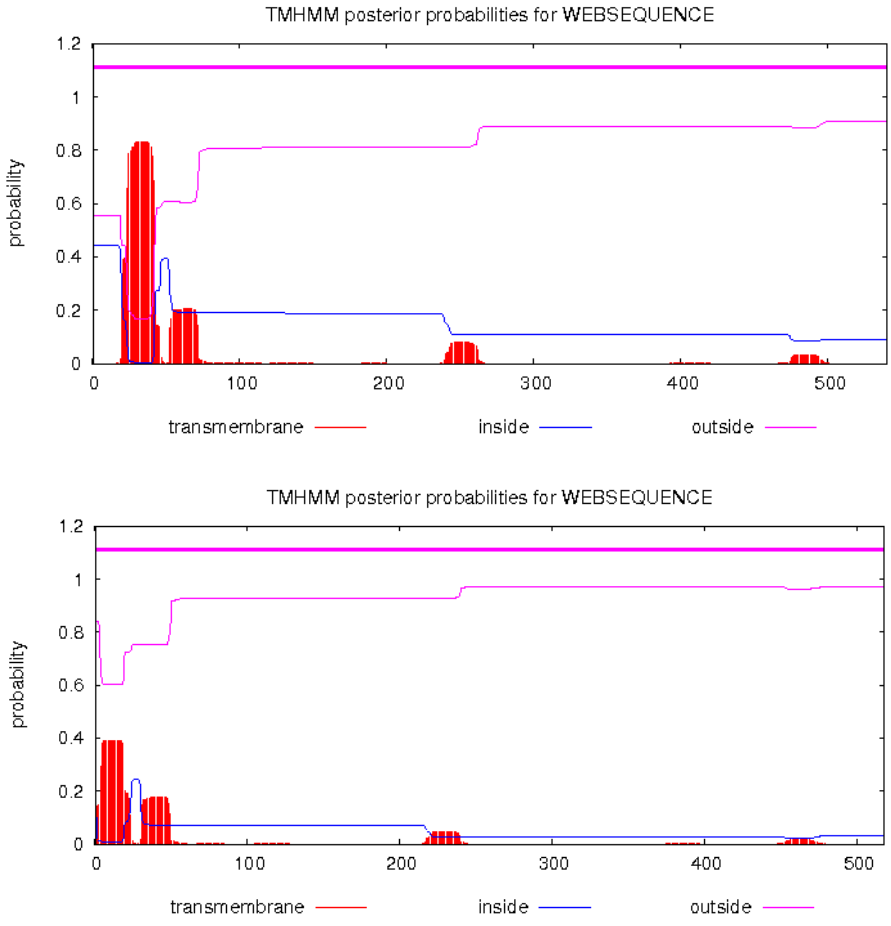


Fig. 2 Amino acid composition of BocCYP79B2s.



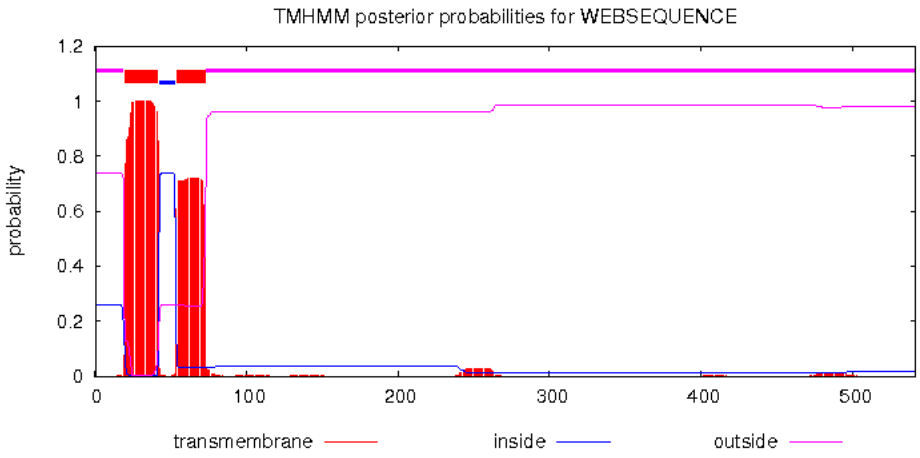


Fig. 3 Prediction of protein transmembrane structure of BocCYP79B2s

3.3. Subcellular localization and conserved domain analysis

Subcellular localization of the *BocCYP79B2.1*, *BocCYP79B2.2* and *BocCYP79B2.3* were all predicted by WoLF PSORT to be in the chloroplast. The analysis using Conserved Domain Database (CDD) demonstrated that the amino acid sequence of all BocCYP79B2 proteins belonged the p450 superfamily and has the conserved domain PLN02971.

3.4. Phylogenetic Tree analysis

A phylogenetic tree was constructed to illustrate the relationship among the CYP79B2 proteins of cabbage and 17 other higher plant species (Fig. 4). A total of three major clusters were identified, BocCYP79B2.1 and BocCYP79B2.3 belong the first cluster, and BocCYP79B2.2 belongs the second cluster. Sequence alignment indicated that the BocCYP79B2.1 protein is more closely related to *Jatropha curcas*, BocCYP79B2.2 protein is more closely related to *Eucalyptus grandis*, and BocCYP79B2.3 protein is more closely related to *Brassica rapa*.

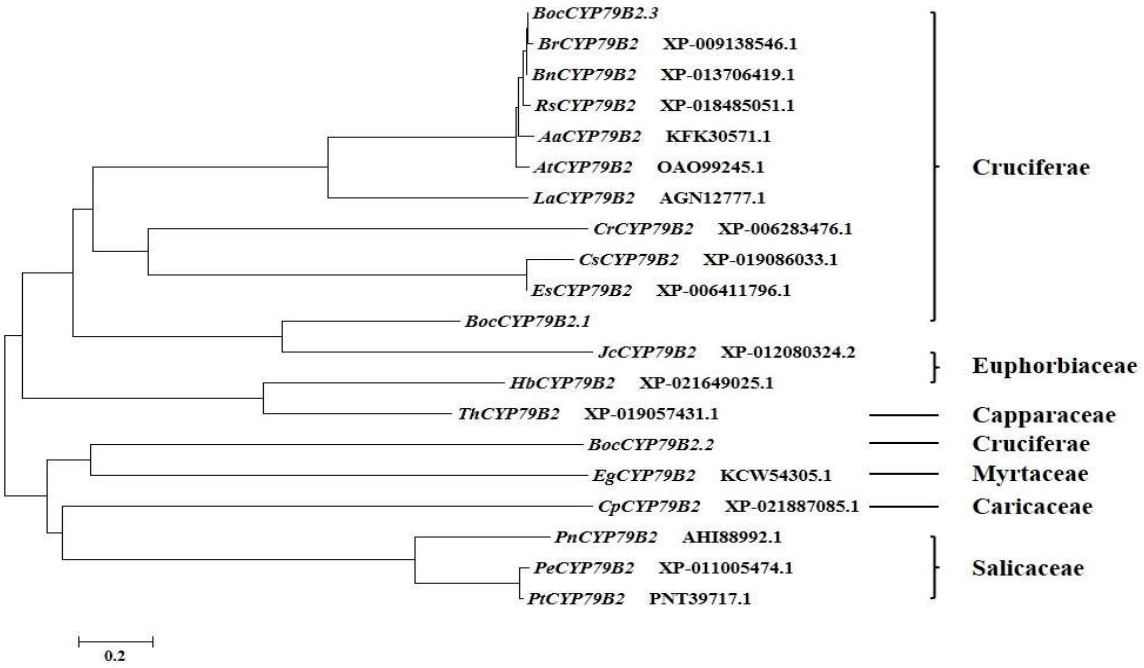


Fig. 4 Phylogenetic tree analysis of BocCYP79B2s and CYP79B2 proteins of other species.

4. Conclusion

The present study analyzed the *BocCYP79B2* gene of cabbage. Zang et al. (2008) transformed the *Arabidopsis CYP79B2* and *CYP83B1* into Chinese cabbage to modulate the indole glucosinolate pathway flux. It was found that overexpression of single *CYP79B2* did not affect the profiles of indole glucosinolates. However, co-expressing *CYP79B2* with *CYP83B1* significantly increased the contents of glucobrassicin, 4-hydroxy glucobrassicin and 4-methoxy glucobrassicin [8]. Previous studies have shown that the CYP79B2 protein is relatively conserved in plants [7]. The findings of the present study show that CYP79B2 from cabbage is highly conserved in plants, similar to that observed in earlier reports. The findings of the present study may serve as a foundation for future studies on the functions of CYP79B2 in glucosinolate metabolism in cabbage.

Acknowledgments

This work was supported by National Natural Science Foundation of China (31500247), key project of Department of Education of Sichuan Province (14ZA0016), and National Student Innovation Training Program (201710626030).

References

[1] M. Wennberg, J. Ekvall, K. Olsson, and M. Nyman, Changes in carbohydrate and glucosinolate composition in white cabbage (*Brassica oleracea* var. *capitata*) during blanching and treatment with acetic acid, Food Chem. 95 (2006) 226-236.

[2] S. Rokayya, C.J. Li, Y. Zhao, Y. Li, and C.H. Sun, Cabbage (*Brassica oleracea* L. var. *capitata*) phytochemicals with antioxidant and anti-inflammatory potential, Asian Pac. J. Cancer Prev. 14 (2014) 6657-6662.

[3] B.G. Hansen, R.E. Kerwin, J.A. Ober, V.M. Lambrix, T. Mitchell-Olds, J. Gershenzon, B.A. Halkier, and D.J. Kliebenstein, A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goiterogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in Arabidopsis, Plant Physiol. 148 (2008) 2096-2108.

[4] X. Yan, and S. Chen. Regulation of plant glucosinolate metabolism, Planta 226 (2007) 1343-1352.

[5] E. Andréasson, L.B. Jørgensen, A.S. Höglund, L. Rask, and J. Meijer, Different myrosinase and idioblast distribution in Arabidopsis and *Brassica napus*, Plant Physiol. 127 (2001) 1750-1763.

[6] N. Benkeblia, Phytonutritional Improvement of Crops, John Wiley & Sons, Inc., Hoboken, 2017, pp. 407-733.

[7] A.K. Hull, R. Vij, and J.L. Celenza, Arabidopsis cytochrome P450s that catalyze the first step of tryptophan-dependent indole-3-acetic acid biosynthesis, P. Natl. Acad. Sci. USA 97 (2000) 2379-2384.

[8] Y.X. Zang, M.H. Lim, B.S. Park, S.B. Hong, and D.H. Kim, Metabolic engineering of indole glucosinolates in Chinese cabbage plants by expression of *Arabidopsis* CYP79B2, CYP79B3, and CYP83B1, Mol. Cells 25 (2008) 231-241.